



# Tissue intrinsic fluorescence recovering by an empirical approach based on the PSO algorithm and its application in type 2 diabetes screening

YUANZHI ZHANG,<sup>1,2,3</sup> HUAYI HOU,<sup>1,3</sup> YANG ZHANG,<sup>1,2</sup> YIKUN WANG,<sup>1,2,\*</sup>  
LING ZHU,<sup>1,2</sup> MEILI DONG,<sup>1,2</sup> AND YONG LIU<sup>1,2</sup>

<sup>1</sup>*Institute of Applied Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Anhui Provincial Engineering Technology Research Center for Biomedical Optical Instrument, Hefei, Anhui 230088, China*

<sup>2</sup>*Wanjiang Center for Development of Emerging Industrial Technology, Tongling, Anhui 244000, China*

<sup>3</sup>*Authors contributed equally to this work*

\*[wky@aiofm.ac.cn](mailto:wky@aiofm.ac.cn)

**Abstract:** In order to reduce the influence of scattering and absorption on tissue fluorescence spectra, after tissue fluorescence and diffuse reflectance in different tissue optical properties were simulated by the Monte Carlo method, a tissue intrinsic fluorescence recovering algorithm making use of diffuse reflectance spectrum was developed. The empirical parameters in the tissue intrinsic fluorescence recovering algorithm were coded as a particle in the solution domain, the classification performance was defined as the fitness, and then a particle swarm optimization (PSO) algorithm was established for empirical parameters optimization. The skin autofluorescence and diffuse reflectance spectra of 327 subjects were collected in Anhui Provincial Hospital. The skin intrinsic autofluorescence spectra were recovered by using the empirical approach and the integration area of the spectra were calculated as fluorescence intensity. Receiver operating characteristic (ROC) analysis for fluorescence intensity was applied to evaluate the classification performance in type 2 diabetes screening. In addition, a support vector machine (SVM) method was implemented to improve the performance of the classification. The results showed that the sensitivity and specificity were 32% and 76% respectively, and the area under the curve was 0.54 before recovering, while the sensitivity and specificity were 72% and 86% respectively, and the area under the curve was 0.86 after recovering. Furthermore, the sensitivity and specificity increased to 83% and 86% respectively when using linear SVM while 84% and 88%, respectively, when using nonlinear SVM. The results indicate that using the tissue fluorescence spectrum recovery algorithm based on PSO can improve the application of tissue fluorescence spectroscopy effectively.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

**OCIS codes:** (170.0170) Medical optics and biotechnology; (260.2510) Fluorescence.

## References and links

1. M. Müller and B. H. Hendriks, "Recovering intrinsic fluorescence by Monte Carlo modeling," *J. Biomed. Opt.* **18**(2), 027009 (2013).
2. M. Y. Berezin and S. Achilefu, "Fluorescence lifetime measurements and biological imaging," *Chem. Rev.* **110**(5), 2641–2684 (2010).
3. A. Kim, M. Khurana, Y. Moriyama, and B. C. Wilson, "Quantification of in vivo fluorescence decoupled from the effects of tissue optical properties using fiber-optic spectroscopy measurements," *J. Biomed. Opt.* **15**(6), 067006 (2010).
4. R. S. Bradley and M. S. Thorniley, "A review of attenuation correction techniques for tissue fluorescence," *J. R. Soc. Interface* **3**(6), 1–13 (2006).
5. Q. Zhang, M. G. Müller, J. Wu, and M. S. Feld, "Turbidity-free fluorescence spectroscopy of biological tissue," *Opt. Lett.* **25**(19), 1451–1453 (2000).
6. S. K. Chang, N. Marín, M. Follen, and R. Richards-Kortum, "Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia," *J. Biomed. Opt.* **11**(2), 024008 (2006).

7. L. Wang, S. L. Jacques, and L. Zheng, "MCML-Monte Carlo modeling of light transport in multi-layered tissues," *Comput. Methods Programs Biomed.* **47**(2), 131–146 (1995).
8. R. Weersink, M. S. Patterson, K. Diamond, S. Silver, and N. Padgett, "Noninvasive measurement of fluorophore concentration in turbid media with a simple fluorescence /reflectance ratio technique," *Appl. Opt.* **40**(34), 6389–6395 (2001).
9. P. A. Valdés, F. Leblond, A. Kim, B. C. Wilson, K. D. Paulsen, and D. W. Roberts, "A spectrally constrained dual-band normalization technique for protoporphyrin IX quantification in fluorescence-guided surgery," *Opt. Lett.* **37**(11), 1817–1819 (2012).
10. N. Nass, B. Bartling, A. Navarrete Santos, R. J. Scheubel, J. Börgermann, R. E. Silber, and A. Simm, "Advanced glycation end products, diabetes and ageing," *Z. Gerontol. Geriatr.* **40**(5), 349–356 (2007).
11. D. C. Bos, W. L. de Ranitz-Greven, and H. W. de Valk, "Advanced glycation end products, measured as skin autofluorescence and diabetes complications: a systematic review," *Diabetes Technol. Ther.* **13**(7), 773–779 (2011).
12. J. Vouillarmet, D. Maucourt-Boulch, P. Michon, and C. Thivolet, "Advanced glycation end products assessed by skin autofluorescence: a new marker of diabetic foot ulceration," *Diabetes Technol. Ther.* **15**(7), 601–605 (2013).
13. E. Hull, M. Ediger, A. Unione, E. Deemer, M. Stroman, and J. Baynes, "Noninvasive, optical detection of diabetes: model studies with porcine skin," *Opt. Express* **12**(19), 4496–4510 (2004).
14. Y. Zhang, L. Pan, Y. Wang, H. Hou, L. Zhu, A. Wang, and Y. Liu, "Calibration for Tissue Fluorescence Spectroscopy System," *Las. Optoelect. Prog.* **52**(8), 081701 (2015).
15. G. M. Palmer and N. Ramanujam, "Monte Carlo-based inverse model for calculating tissue optical properties. Part I: Theory and validation on synthetic phantoms," *Appl. Opt.* **45**(5), 1062–1071 (2006).
16. X. Zhong, X. Wen, and D. Zhu, "Lookup-table-based inverse model for human skin reflectance spectroscopy: two-layered Monte Carlo simulations and experiments," *Opt. Express* **22**(2), 1852–1864 (2014).
17. J. Kennedy, "Particle swarm optimization," *Encyclopedia of Machine Learning* (Springer US, 2011), pp. 760–766.
18. R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm Intell.* **1**(1), 33–57 (2007).
19. J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.* **9**(3), 293–300 (1999).
20. E. Salomatina, B. Jiang, J. Novak, and A. N. Yaroslavsky, "Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range," *J. Biomed. Opt.* **11**(6), 064026 (2006).
21. A. J. Smit, J. M. Smit, G. J. Botterblom, and D. J. Mulder, "Skin autofluorescence based decision tree in detection of impaired glucose tolerance and diabetes," *PLoS One* **8**(6), e65592 (2013).
22. D. R. Jesudason, K. Dunstan, D. Leong, and G. A. Wittert, "Macrovascular risk and diagnostic criteria for type 2 diabetes: implications for the use of FPG and HbA1c for cost-effective screening," *Diabetes Care* **26**(2), 485–490 (2003).
23. J. D. Maynard, M. Rohrscheib, J. F. Way, C. M. Nguyen, and M. N. Ediger, "Noninvasive type 2 diabetes screening: superior sensitivity to fasting plasma glucose and A1C," *Diabetes Care* **30**(5), 1120–1124 (2007).
24. Y. Zhang, L. Zhu, Y. Wang, L. Zhang, S. Ye, Y. Liu, and G. Zhang, "Classification of skin autofluorescence spectrum using support vector machine in type 2 diabetes screening," *J. Innov. Opt. Health Sci.* **6**(04), 1350036 (2013).
25. L. Zhu, Y. Wang, S. Ye, H. Wang, Y. Zhang, F. Li, C. Ye, L. Zhang, and Y. Liu, "Investigation on noninvasive and rapid diabetes screening using skin autofluorescence," *Chinese Journal of Diabetes Mellitus* **5**(005), 293–296 (2013).

## 1. Introduction

Fluorescence, the re-emission of light by fluorophore that has absorbed a shorter wavelength light, is widely used to investigate biological tissues [1, 2]. The shape and intensity of the fluorescence spectrum contain valuable information on the identity and density of fluorophore in tissue. However, in samples with high scattering or absorption, such as tissue, both the shape and intensity of measured fluorescence spectrum can be heavily distorted, and then resulting in the measured fluorescence spectrum may not be proportional to fluorophore concentration [3]. Untangling the effects of this attenuation on the measured fluorescence spectrum is necessary for truly quantitative analysis.

Several methods have been reported to disentangle the effects of absorption and scattering from the measured fluorescence spectrum to recover the intrinsic fluorescence spectrum. Bradley et al. reviewed over 50 different publications that addressed the recovery of intrinsic fluorescence spectrum [4]. These studies have used theoretical methods based on physical models of light tissue interactions, including analytical approaches based on diffusion theory [5, 6] as well as computational techniques such as Monte Carlo simulations of photon transport in turbid media [7] or simple empirical approaches [8, 9]. Although promising

results have been obtained with various methods, thus far, all available methods have their limitation.

Diffusion theory is an approximation to the radiative transfer equation and it describes photons transporting in absorbing and scattering media. Thus it can correct the measured fluorescence spectrum based on the tissue absorption and scattering properties to extract the intrinsic fluorescence spectrum analytically. However, the diffusion approximation is only valid when the absorption coefficient of the medium is at least an order of magnitude lower than scattering, and the distances between sources and detectors are much greater than the mean free path of diffusing photons in that medium.

The Monte Carlo method is stochastic in nature and is not limited to the diffusion regime. It can therefore be used to model light transport over the entire UV-visible-NIR range. However, Monte Carlo simulations are computationally time-consuming and have historically not been very convenient to use as inverse models. Empirical approach, including subtraction techniques and ratio techniques, always employ the measurements of the reflected illumination intensity in order to reduce the effects of tissue optical properties on fluorescence. In subtraction techniques, attenuation in fluorescence due to tissue optical properties is compensated for by subtracting a proportion of the corresponding change in the reflected excitation light intensity from the fluorescence intensity. Subsequently, investigators argue that it would be more appropriate to utilize a ratio of fluorescence intensity to diffuse reflectance intensity rather than an algebraic subtraction. However, there has been no clear consensus on how to realize the ratio techniques and how to determine the form and the parameters of ratio equation. In addition, as the same to subtraction technique, ratio techniques also lack the model foundation.

Advanced glycation end products (AGE) are a complex and heterogeneous group of compounds that have been implicated in diabetes related complications [10]. Skin autofluorescence is related to the accumulation of AGE and has a potential to provide prognostic information of diabetes and its related complications [11, 12]. While the skin autofluorescence is heavily distorted by the tissue scattering and absorption, resulting in the measured autofluorescence not be proportional to AGE accumulation. In this paper, we simulated the tissue fluorescence and diffuse reflectance at both the excitation and the emission wavelength in different tissue optical properties by Monte Carlo method and introduced a tissue intrinsic fluorescence recovering algorithm [13], which made use of a diffuse reflectance measurement taken at the same location. In order to ensure the recovering performance, a particle swarm optimization (PSO) algorithm was introduced to accomplish the empirical parameters optimization. Subsequently, we introduced this recovering algorithm to untangle the effects of tissue attenuation on the measured fluorescence spectrum and used support vector machine (SVM) to classify the recovered intrinsic fluorescence for type 2 diabetes screening.

## 2. Materials and methods

### 2.1 Subjects

A total of 346 subjects (152 male, 194 female) were recruited in Anhui Provincial Hospital in 2014. Meanwhile, 19 subjects were excluded according to the exclusion criteria: 1) subjects unwilling to comply with test specifications; 2) the site of the skin measured has scar, moss-like sclerosis spots, vitiligo, deformity and infectious skin disease. At last, 327 subjects (143 male, 184 female) were included. The mean age was  $(47 \pm 17)$  years, mean height was  $(164 \pm 8)$  cm, mean weight was  $(64.7 \pm 11.4)$  kg and mean BMI was  $(23.6 \pm 3.7)$  kg/m<sup>2</sup>. Oral glucose tolerance test (OGTT) was carried out for all subjects. And then, all subjects were divided into diabetes mellitus (DM) group and non-diabetic mellitus (NDM) group based on OGTT-2h-value (If OGTT-2h-value was greater than 11.1/L, determine the subject with diabetes, otherwise, determine non-diabetic). Ultimately, there were 208 subjects in DM group, and 119 subjects in NDM group. Additionally, skin autofluorescence was measured for both DM

and NDM groups, everyone was measured three times and the average was taken as the final value. The trial was approved by the Medical Ethics Committee of Anhui Provincial Hospital, and informed consent form was obtained from the subjects.

## 2.2 Instrumentation

The optical system, as depicted in Fig. 1(a), has been described particularly in our previous work [14]. In short, it consisted of an ultraviolet light source as excitation source, a broadband light source, a trifurcated fiber-optic probe, and a compact charge-coupled device (CCD) spectrometer. In addition, there were other parts such as power supply units, control units, optical and mechanical parts.

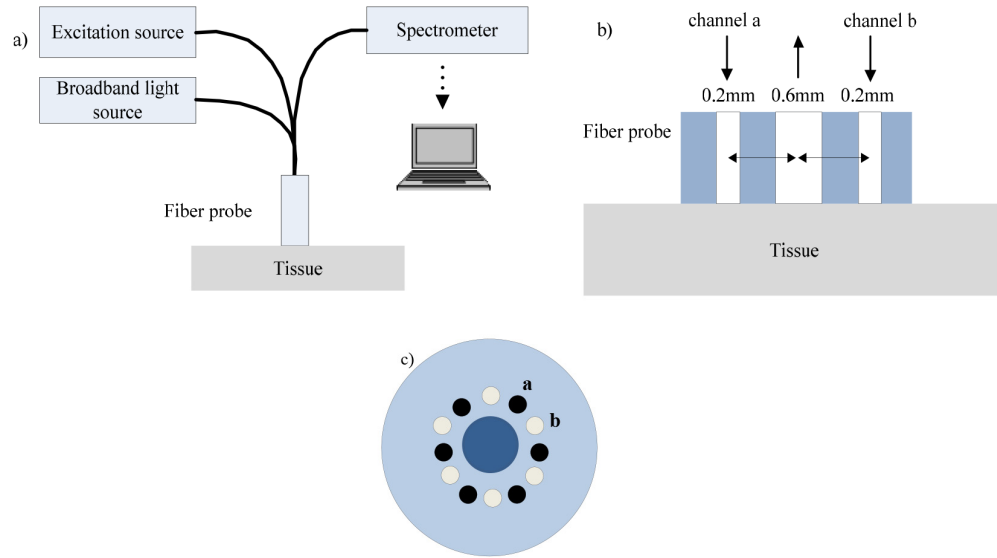


Fig. 1. Schematic of the optical system with a fiber-optic probe.

The above trifurcated fiber-optic probe consisted of two channels, as depicted in Fig. 1(b), identified as a and b. The channel a, which connected to the ultraviolet light source, was designed for fluorescence measurements. The channel b, which connected to the broadband light source, was used for reflectance measurements. The illumination fibers of channels a and b were arranged in a concentric circle at the probe distal end as Fig. 1(c). Moreover, the numerical aperture of all fibers was 0.22, and the core diameters of illumination and collection fibers were 200 and 600  $\mu\text{m}$ , respectively.

Convolution was used to integrate over the illumination and collection fibers in order to determine the probability that a photon traveling a fixed distance and would be collected [15, 16]. Considering current probes's geometry, the probability of collection of a photon traveling a net distance  $r_t$  between the points of entering and exiting the medium is given by:

$$p = 6 \cdot \frac{1}{\pi^2 r_i^2} \int_{lb}^{ub} (s-x) \cos^{-1} \left[ \frac{s^2 + (s-x)^2 - r_i^2}{2(s-x)s} \right] \cos^{-1} \left[ \frac{r_t^2 + (s-x)^2 - r_c^2}{2(s-x)r_t} \right] dx \quad (1)$$

Where  $ub = \min(r_i, s - r_t + r_c)$ ,  $lb = \max(-r_i, s - r_t - r_c)$ , the factor of 6 accounts for the six illumination fibers surrounding the collection fiber,  $r_i$  means the radius of the illumination fiber,  $r_c$  means the radius of the collection fiber,  $s$  means the distance between the center of illumination and collection fibers, which is 500  $\mu\text{m}$ ,  $r_t$  is equal to  $s$ .

Finally, the excitation wavelength was 370 nm, and the corresponding emission wavelength was 480 nm. The spectrometer was USB4000-UV-VIS fiber optic spectrometer of Ocean Optics and the range of integration times was 3.8 ms ~10 s.

Figure 2 showed the flow chart of tissue fluorescence and diffusion reflectance collection. During the fluorescence measurement, excitation source was activated; while in diffusion reflectance spectra measurement, the broadband light source was activated. Ultimately, the fluorescence and diffuse reflectance spectra were transmitted to the computer.

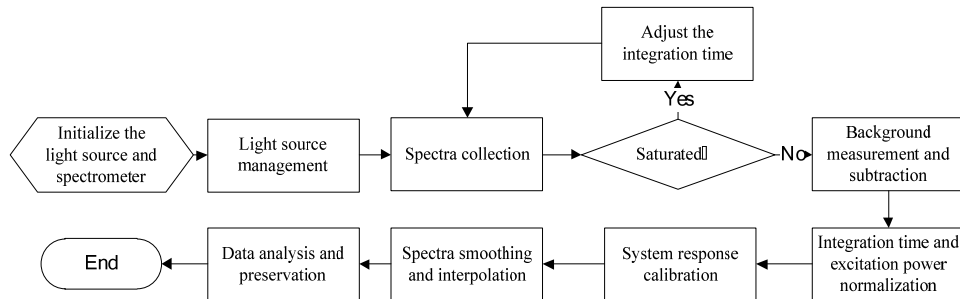


Fig. 2. Flow chart of spectra collection.

### 2.3 Recovery algorithm

Tissue diffuse reflectance spectrum was measured at the same site with the fluorescence, and it could be employed to recover the distortion of fluorescence due to tissue absorption and scattering. To extract the intrinsic fluorescence from the extrinsic measured fluorescence, an empirical recovery algorithm based on diffuse reflectance spectrum was established. The central mission was to determine the equation form and the corresponding parameters.

$$f(\lambda_m) = \text{function}(F(\lambda_x), R(\lambda_x), R(\lambda_m)) \quad (2)$$

The recovery algorithm described in Eq. (2) was based on the assumption that the optical absorption at the excitation wavelength  $\lambda_x$  was high relative to that at emission wavelength  $\lambda_m$ . This was generally true in tissue if the excitation wavelength was in the UV to the blue end of the visible spectrum. As a result, most fluorophore absorption events occurred close to the source fiber. The migration paths of the fluorescence emission photons at  $\lambda_m$  were then approximated by those of the reflectance photons at  $\lambda_m$ , emitted and collected using the same fiber optic geometry. The tissue optical model for fluorescence and diffuse reflectance collection were shown in Fig. 3.

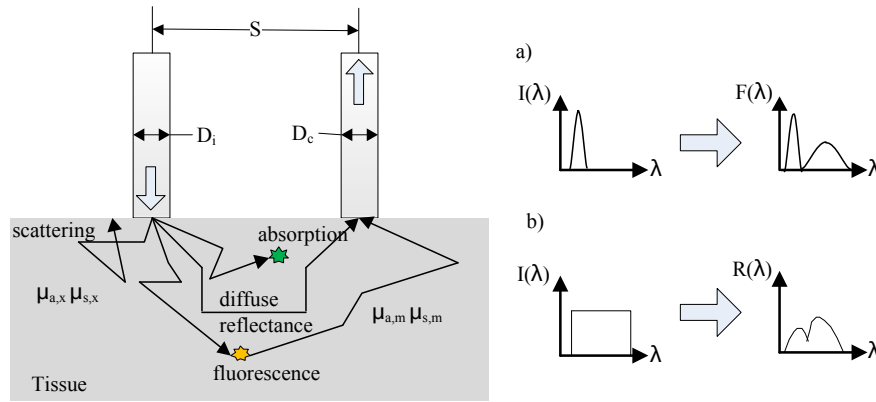


Fig. 3. Tissue optical model (a) for fluorescence measurement; (b) for diffuse reflectance spectra measurement.

## 2.4 Particle swarm optimization (PSO) algorithm

The particle swarm is a population based stochastic algorithm for optimization which is based on social psychological principles [17, 18]. In PSO algorithm, the particles are placed in the search space, and each evaluates the objective function at its current location. Then each particle determines its movement through the search space by combining some aspect of the history of its own current and best locations with those of one or more members of the swarm, with some random perturbations. The next iteration takes place after all particles are moved. Eventually the swarm as a whole, like a flock of birds collectively foraging for food, is likely to move close to an optimum of the fitness function.

Each individual in the particle swarm is composed of three D-dimensional vectors, where  $D$  is the dimensionality of the search space. These are the current position  $x_i$ , the velocity  $v_i$ , and the previous best position  $p_i$  for each individual. In addition, for the particle swarm, there is the global best position  $p_g$ . The process for implementing PSO is shown in Algorithm 1. Figure 4 showed the flow chart of PSO algorithm.

### Algorithm 1

1. Initialize a population array of particles with random positions and velocities on D-dimensions in the search space.
2. Loop.
3. For each particle, evaluate the desired optimization fitness function in D variables.
4. Compare particle's fitness evaluation with its  $pbest_i$ . If current value is better than  $pbest_i$ , then set  $pbest_i$  equal to the current value, and  $p_i$  equal to the current location  $x_i$  in D-dimensional space.
5. Identify the particle in the neighborhood with the best success so far, and assign its index to the variable  $g$ .
6. Change the velocity and position of the particle according to the following equation

$$v_{id}(t+1) = \omega v_{id}(t) + \eta_1 rand(p_{id} - x_{id}(t)) + \eta_2 rand(p_{gd} - x_{id}(t)) \quad (3)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (4)$$



where  $\omega$  is inertia weight;  $v_{id}$  is particle velocity;  $p_{id}$  is optimal individual particle;  $p_{gd}$  is optimal global particle;  $\eta_1$ ,  $\eta_2$  are learning factors;  $x_{id}$  is particle location and  $rand$  is random number between 0 and 1.

7. If a criterion is met (usually a sufficiently good fitness or a maximum number of iterations), exit loop.
8. End loop.

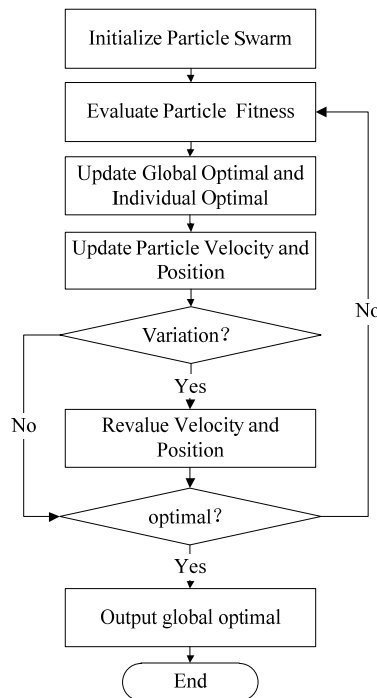


Fig. 4. Flow chart of PSO algorithm.

## 2.5 Support vector machine

Support vector machine (SVM) is a machine-learning method based on the principle of structural risk minimization and originally developed by Vapnik and Burges [19]. It has many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition. The main mechanism of SVM is to hunt an optimal separating hyper plane that meets the classification requirements. The plane should ensure the required classification accuracy, as well as make the classification interval maximum. In theory, SVM can achieve the optimal classification for linearly separable problems. For nonlinear separable problems, they are firstly mapped into a high dimensional linearly separable space through a nonlinear mapping, and then traded as linearly separable problems. The kernel functions used in constructing nonlinear SVM classifiers are the polynomial function, the radial basic function (RBF), and the neural network function, etc. One of the most used kernel functions in the reported work is the RBF kernel defined as:

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \quad (5)$$

During the classification using SVM, two aspects should be considered: how to choose the optimal input feature subset for SVM, and how to set the best kernel parameters. These two aspects are crucial, because the feature subset choice influences the appropriate kernel

parameters and vice versa. Therefore, obtaining the optimal feature subset and SVM parameters must occur simultaneously. In this paper, skin autofluorescence was chosen as input feature. The parameters that should be optimized include penalty parameter  $C$  and the kernel function parameters such as parameter  $\gamma$  for the RBF kernel. To design a SVM, one must choose a kernel function, set the kernel parameters and determine the penalty parameter. Penalty parameter represents the compromise on training error and generalization ability. Cross-validation and grid-optimization methods are alternative to determine the kernel function and the optimal parameters. In this work, linear SVM and nonlinear SVM based on RBF kernel were used to classify the skin autofluorescence of control subjects and diabetes.

### 3. Results and discussion

#### 3.1 Tissue intrinsic fluorescence recovery algorithm

Monte Carlo method was used to simulate the tissue fluorescence and diffuse reflectance in different tissue optical properties. During the simulation, we made some assumptions. Firstly, the illumination fiber and the collection fiber were set to the same type of 200/220 $\mu\text{m}$  fiber with a numerical aperture (NA) of 0.22, and the separation between the centers of the illumination and the collection fibers was 0.3 mm. Secondly, the size of tissue was set to 4 mm  $\times$  4 mm, and the thickness was 1mm. For the excitation wavelength, we set tissue absorption coefficient ranging from 1 to 15  $\text{cm}^{-1}$  and reduced scattering coefficient to 5.1 $\text{cm}^{-1}$ ; for the emission wavelength, we set tissue absorption coefficient is 1.1  $\text{cm}^{-1}$ , and reduced scattering coefficient ranging from 10 to 150  $\text{cm}^{-1}$ . The refractive index of tissue was 1.37, the anisotropic properties of tissue was 0.9 [20]. Because the scattering coefficient at excitation wavelength  $u_{s,x}$  and absorption coefficient at emission wavelength  $u_{s,m}$  were constants, scattering coefficient meant the reduced scattering coefficient at emission wavelength  $u_{s,m}$  and absorption coefficient meant the absorption coefficient at excitation wavelength  $u_{a,x}$  in the following description.

The simulation results were shown in Fig. 5, among them,  $k_x$  and  $k_m$  were semi-empirical parameters. Figure 5(a) showed tissue fluorescence  $F(\lambda_m)$ , diffuse reflectance at excitation wavelength  $R_x$  and diffuse reflectance at emission wavelength  $R_m$  in different scattering coefficient, when absorption coefficient  $u_{a,x}$  was 5  $\text{cm}^{-1}$ . In Fig. 5(a),  $R_x$  remained constant, while the  $F(\lambda_m)$  and  $R_m$  increased as the tissue scattering coefficient rising, and  $R_m$  increased significantly faster than  $F(\lambda_m)$ .

Figure 5(b) showed  $F(\lambda_m)$ ,  $R_x$  and  $R_m$  in different absorption coefficient, when scattering coefficient  $u_{s,m}$  was 50  $\text{cm}^{-1}$ . As showed in this Fig.,  $R_m$  remained unchanged, while the  $F(\lambda_m)$  and  $R_x$  decreased as the tissue scattering coefficient rising, and  $R_x$  decreased few faster than  $F(\lambda_m)$ .

Figure 5(c) showed  $F(\lambda_m)$ , the ratio of  $F(\lambda_m)$  to  $R_m$ , the ratio of  $F(\lambda_m)$  to  $(R_m)^{k_m}$  in different scattering coefficient, when absorption coefficient  $u_{a,x}$  was 5  $\text{cm}^{-1}$ . In this Fig., the ratio of  $F(\lambda_m)$  to  $R_m$  decreased as the tissue scattering coefficient rising, while  $F(\lambda_m)$ , the ratio of  $F(\lambda_m)$  to  $(R_m)^{k_m}$  increased as the tissue scattering coefficient rising, and the change of the ratio of  $F(\lambda_m)$  to  $(R_m)^{k_m}$  was smaller than  $F(\lambda_m)$ .



Figure 5(d) showed  $F(\lambda_m)$ , the ratio of  $F(\lambda_m)$  to  $R_x$ , the ratio of  $F(\lambda_m)$  to  $(R_x)^{k_x}$  in different absorption coefficient, when scattering coefficient  $u_{s,m}$  was  $50 \text{ cm}^{-1}$ . In this Fig.,  $F(\lambda_m)$  decreased as the tissue absorption coefficient rising, while the ratio of  $F(\lambda_m)$  to  $R_x$ , the ratio of  $F(\lambda_m)$  to  $(R_x)^{k_x}$  varying in a certain range as the tissue scattering coefficient rising, and the change of the ratio of  $F(\lambda_m)$  to  $(R_x)^{k_x}$  was smaller than the ratio of  $F(\lambda_m)$  to  $R_x$ .

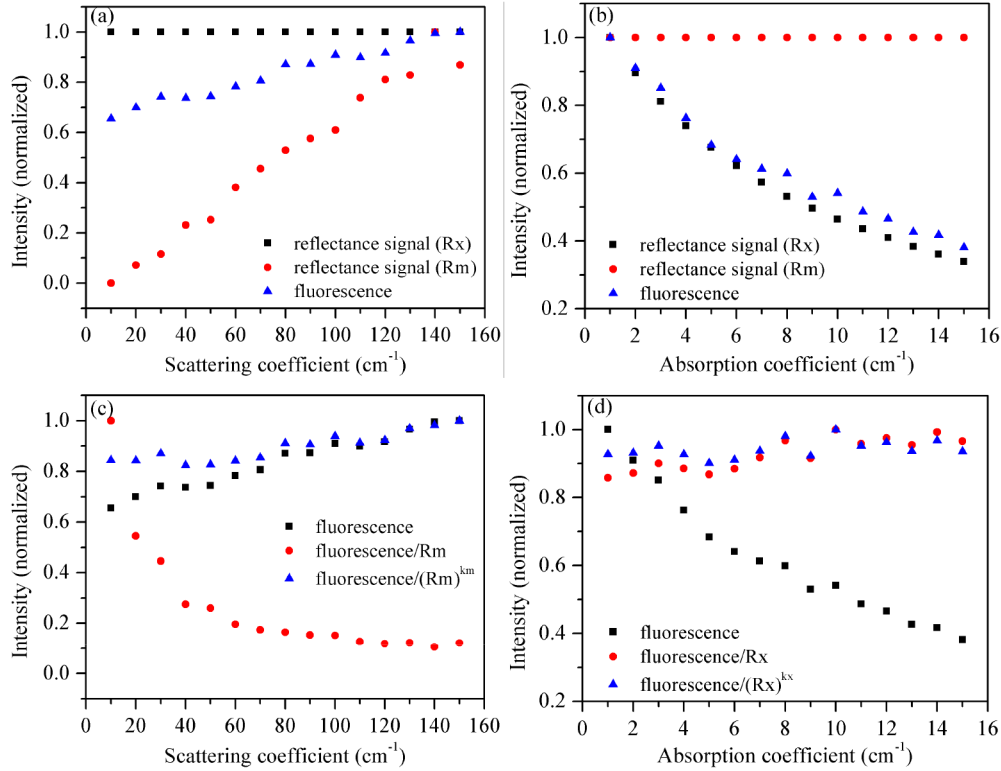


Fig. 5. Results of Monte Carlo modeling (a), (c):  $u_{a,x} = 5 \text{ cm}^{-1}$ ; (b), (d):  $u_{s,m} = 50 \text{ cm}^{-1}$

It could be seen that most of the changes in fluorescence magnitude due to tissue absorption at the excitation wavelength and tissue scattering at the emission wavelength could be corrected by further dividing the raw fluorescence and by an empirical power function of diffuse reflectance at excitation wavelength and diffuse reflectance at emission wavelength respectively. Furthermore, we can extend singleness emission wavelength to emission spectrum specific tissue. So an empirical approach, described as Eq. (6), was introduced to recover the intrinsic fluorescence [13].

$$f(\lambda_m) = \frac{F(\lambda_m)}{R(\lambda_x)^{k_x} R(\lambda_m)^{k_m}} \quad (6)$$

where  $F(\lambda_m)$  was the measured fluorescence spectrum,  $R(\lambda_x)$  and  $R(\lambda_m)$  were the spectrally integrated reflectance signals at the excitation wavelength and the emission wavelength, respectively. Both  $F(\lambda_m)$ ,  $R(\lambda_x)$  and  $R(\lambda_m)$  were calculated use the raw

measured spectrum divided the collection probability. The  $k_x$  and  $k_m$  were semi-empirical parameters ranging from 0 ~1.

### 3.2 Recovery parameters optimization

Diabetes mellitus (DM) and impaired glucose tolerance (IGT) detection are conventionally based on glycemic criteria. Skin autofluorescence (SAF) is a noninvasive proxy of tissue accumulation of advanced glycation end products (AGE) which is considered to be a carrier of glycometabolic memory [21]. While the detected skin autofluorescence spectrum is heavily distorted due to emitted fluorescence being absorbed and scattered within the sample, and furthermore, the propagation of the excitation light is affected by the background absorption and scattering, they result in that the detected autofluorescence may not be proportional to fluorophore concentration. In order to disentangle the effects of absorption and scattering from the measured autofluorescence spectrum, the empirical algorithm described in section 3.1 was introduced, and then we used the PSO algorithm to optimize the recovery parameters. Table 1 showed the PSO parameters for recovery parameters optimization.

**Table 1. Description of PSO parameters for recovery parameters optimization**

Character	Description
Particle	Parameter vector, $\{k_x, k_m\}$
Particle swarm	The set of parameter vectors, $\{k_x, k_m\}_1, \{k_x, k_m\}_2, \dots, \{k_x, k_m\}_n$
Velocity of particle	The update velocity of parameter vector, $\{v_x, v_m\}$
Position of particle	The value of parameter vector
Fitness of particle	The classification performance of fluorescence
Individual optimum	The best position of a particle during <u>iteration</u>
Global optimum	The best position of a particle swarm during <u>iteration</u>
Variation condition	Assign values to particles whom fitness distribute in the last 10 percent
Optimum condition	Reach convergence of fitness or the maximum iteration times

During the parameters optimization, recovered skin autofluorescence integrated intensity from 400 nm to 600 nm of both DM and NDM group subjects was analyzed by receiver operating characteristic (ROC), and the fitness of particle was defined as the area under the ROC curve (AUC); population size was set to 20; and the maximum number of iterations was set to 50. Figure 6 showed the optimization results of PSO algorithm. Figure 6(a) showed that the recovery parameters  $k_x$  and  $k_m$  tended to be stable after 17 times iteration, and their values were 0.61 and 0.43 respectively. In the corresponding, the best fitness of particle was 0.86 in Fig. 6(b).

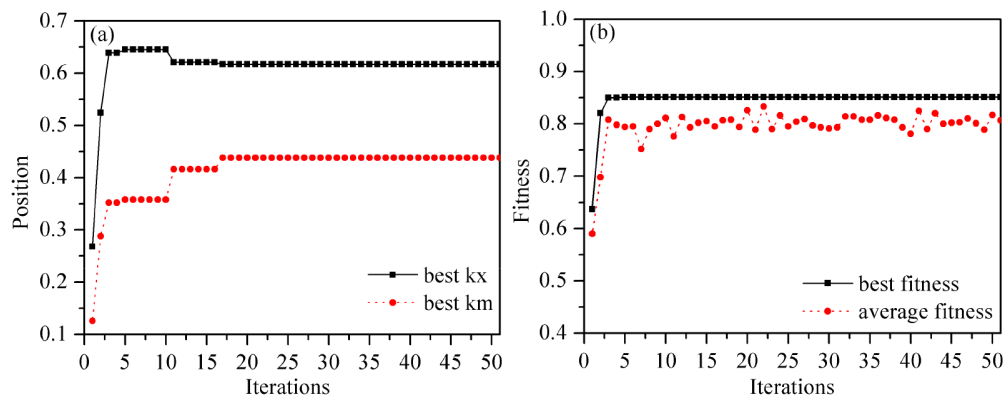


Fig. 6. Optimization results by PSO algorithm (a): particle position; (b): fitness during iteration.

The ROC curve of fluorescence for type 2 diabetes screening under parameter  $\{0.61, 0.43\}$  and four others sets were shown in Fig. 7 and Table 2. Parameters  $\{1, 1\}$ ,  $\{0, 0\}$ ,  $\{1, 0\}$ ,  $\{0, 1\}$  showed  $f(\lambda_m)$  equal to  $F(\lambda_m)/(R(\lambda_x)*R(\lambda_m))$ ,  $F(\lambda_m)$ ,  $F(\lambda_m)/R(\lambda_x)$ ,  $F(\lambda_m)/R(\lambda_m)$ , respectively. For the parameter  $\{0.61, 0.43\}$ , the area under ROC curve was 0.86, significantly higher than the value of 0.54 before recovering. The sensitivity and specificity were 72% and 86% respectively, significantly better than the values of 32% and 76% before recovering.

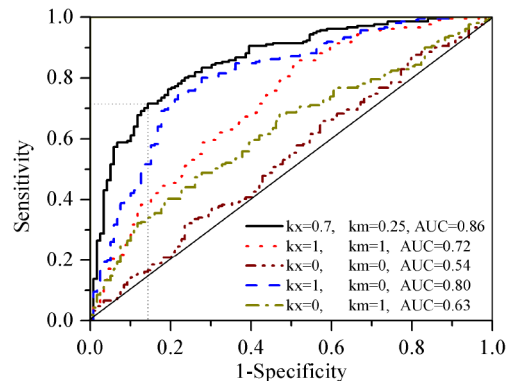


Fig. 7. ROC curves quantifying the performance of five sets recovery parameters.

Figure 8 showed the raw measured fluorescence spectrum and the recovered intrinsic fluorescence spectrum of both DM and NDM group subjects. The red solid and dotted line were the raw measured fluorescence spectrum and the recovered intrinsic fluorescence spectrum of DM group, respectively; Black solid and dotted line were the raw measured fluorescence spectrum and the recovered intrinsic fluorescence spectrum of NDM group, respectively. The raw measured fluorescence spectrum intensity of DM group was significantly higher than the NDM group. Both DM and NDM group, the measured fluorescence spectrum existed peaks at 450 nm, 480 nm and 510 nm. For the measured fluorescence of DM group, the peak at 510 nm was significantly higher than the peak at 480 nm, while for the measured fluorescence of NDM group, the peak at 510 nm was equal to the peak at 480 nm. The measured fluorescence of both DM and NDM group exist an obvious difference in curve shape.

Table 2. Quantifying the performance of five sets recovery parameters

	Sensitivity (%)	Specificity(%)	AUC
$k_x = 0.61, k_m = 0.43$	72	86	0.86
$k_x = 1, k_m = 1$	86	49	0.72
$k_x = 0, k_m = 0$	32	76	0.54
$k_x = 1, k_m = 0$	75	78	0.80
$k_x = 0, k_m = 1$	40	81	0.63
95% CI	31.72~90.28	56.03~91.97	0.551~0.870
$t(\chi^2)$	5.785 <sup>a</sup>	11.432 <sup>a</sup>	12.360 <sup>a</sup>

Notes: CI: confidence intervals; <sup>a</sup> $P < 0.01$ .

While during intrinsic fluorescence recovering, the fluorescence intensity of DM group decreased and the fluorescence shape of DM group was distorted; the fluorescence intensity of NDM group increased and the fluorescence shape of NDM group was distorted. After recovered, the fluorescence spectrum intensity of DM group was still significantly higher than NDM group, but the shape was very similar. For the recovered intrinsic fluorescence spectra of both DM and NDM group, the fluorescence emission peak at 450nm changed to 440 nm. The main emission peak changed to 480 nm, which was more consistent with matrix organization fluorescence spectrum stimulated emission of radiation.

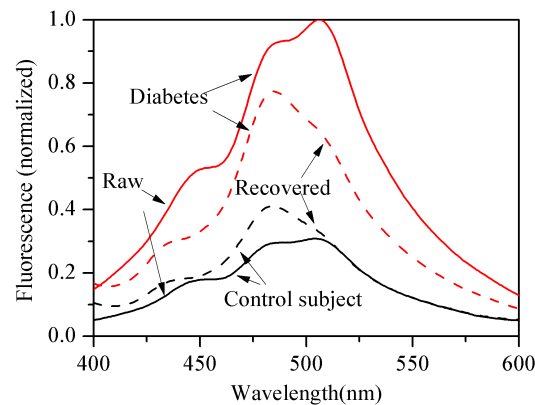


Fig. 8. Skin autofluorescence spectra (measured extrinsic and recovered intrinsic) of DM and NDM group.

### 3.3 Fluorescence classification for type 2 diabetes screening

The dimension of spectrum was generally high. For instance, the tissue intrinsic autofluorescence spectrum used in this study had 200 dimensions (from 400 nm to 600 nm). The high dimension of data space may cause complexity in implementation of the SVM algorithm because computation of all the inner products between the sample and support vectors in a high-dimensional feature space was complicated and time-consuming. In order to simplify the implementation of the SVM algorithm and improve its performance, principal component analysis (PCA) method was introduced to reduce the dimensions of the tissue intrinsic autofluorescence.

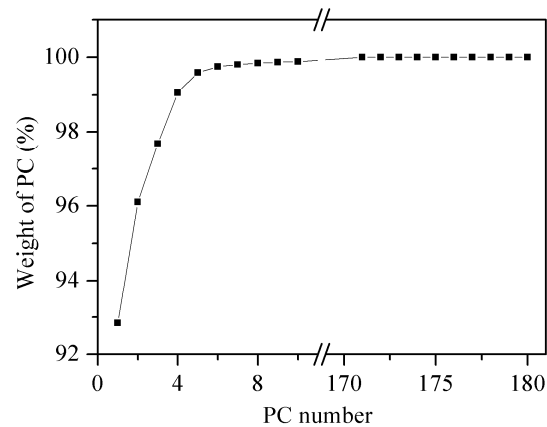


Fig. 9. Contributions of principal components to the total variance of tissue intrinsic fluorescence.

When PCA was used to process autofluorescence spectra, it transformed wavelengths, the original spectral variables, into a set of PC spectra. Each original spectrum was a combination of the PC loading spectra that were orthogonal to each other. The PCs with negligible contributions to the variance of the data set were eliminated. The dimensions of the data set for developing the diagnostic algorithm can be significantly reduced without losing useful information. Figure 9 showed the contribution of PC to the variance of the 327 autofluorescence spectra. The PCs were calculated with a MATLAB based PCA program. As shown in the Fig., the first PC accounted for 92.8% of the total variance, the first two PCs accounted for 96.1% and the first five PCs accounted for 99.5%.

Table 3. Results of classification of autofluorescence with different algorithms

Algorithm	sensitivity	Specificity	Accuracy
Linear SVM	0.83	0.86	82.9
Nonlinear SVM (RBF)	0.84	0.88	84.1

Table 4. Comparison of FPG, HbA1c and noninvasive method

Paper	Subject(DM)	Diagnose criteria OGTT 2hPG	Method	sensitivity	Specificity	AUC
Jesudason [22]	505(177)	$\geq 11.1$ mmol/L	FPG	74.1	94.5	0.91
			HbA1c	80.0	86.3	0.90
Maynard [23]	351(84)	$\geq 140$ mg/dl	FPG	58.0	77.4	0.72
			HbA1c	63.8	77.4	0.79
			Noninvasive	74.7	77.4	0.80
Zhang [24]	203(63)	$\geq 11.1$ mmol/L	FPG	77.8	77.5	-
			Noninvasive	69.6	95.7	-
Zhu [25]	201(63)	$\geq 11.1$ mmol/L	FPG	65.1	86.2	0.83
			Noninvasive	77.8	86.2	0.92
This paper	327(208)	$\geq 11.1$ mmol/L	Noninvasive	84.0	88.0	-

In the process of SVM learning and testing, all of the tissue intrinsic fluorescence spectra were analyzed by PCA method and the first five PCs were chosen as the input vectors. The 327 samples were randomly divided into training set (75%) and testing set (25%). The training set was used to cross validation and then optimize the kernel parameters; the testing

set was used to evaluate the performance of the SVM algorithms. Table 3 showed the results of the classification of the first five PCs. The results demonstrated that both the linear SVM algorithm and nonlinear SVM has a good performance in spectra classification for diabetes screening.

#### 4. Conclusion

Tissue fluorescence spectroscopy is highly sensitive to the microenvironment inside the tissue, and has broad application prospects in cancer tissue detection, photodynamic therapy and other fields.

In this paper, tissue fluorescence and diffuse reflectance in different tissue optical properties were simulated by Monte Carlo method and tissue intrinsic fluorescence recovering algorithm, making use of a diffuse reflectance measurement taken at the same location, was established. The empirical parameters in tissue intrinsic fluorescence recovering algorithm were coded as a particle in solution domain, the classification performance was defined as the fitness, and then a particle swarm optimization (PSO) algorithm was established for empirical parameters optimization.

The tissue intrinsic fluorescence recovering algorithm was used in skin autofluorescence recovering, and then the recovered intrinsic fluorescence was analyzed for diabetes screening combining with SVM. The comparison results of FPG, HbA1c and noninvasive method were shown in Table 4. The first set of data showed the performance of FPG and HbA1c for type 2 diabetes screening introduced by American Diabetes Association. The second to fourth sets of data showed the comparison of FPG, HbA1c and noninvasive method introduced by different groups. While the last set data showed the performance of noninvasive method for type 2 diabetes screening in this paper. From Table 4, noninvasive method based on skin fluorescence showed a better performance than FPG and HbA1c in diabetes screening. And comparing to the previous work, tissue intrinsic fluorescence recovering by empirical approach based on PSO algorithm in this paper can improve the effect of diabetes screening based on fluorescence spectrometry, has an obvious progress.

#### Funding

National Science and Technology Infrastructure Program (2015BAI01B04); Research Project for Practice Development of National TCM Clinical Research Bases (2015D05, supporting institution: Anhui Province Hospital of Chinese Medicine); Science and Technology Service Network Initiative of Chinese Academy of Sciences (KFJ-SW-STS-161); Science and Technology Major Project of Anhui Province (15czz02019).

#### Disclosures

The authors declare that there are no conflicts of interest related to this article.